A Generalization of Permanents and Determinants

D. Vere-Jones
Victoria University of Wellington
Wellington, New Zealand

Submitted by Marvin D. Marcus

ABSTRACT

A method is developed for expanding arbitrary powers of the characteristic polynomial of a matrix. The coefficients are expressed in terms of matrix functions generalizing those of the permanent and determinant. The expansions have applications in probability theory.

1. INTRODUCTION AND STATEMENT OF RESULT

In an earlier note [4] I drew attention to the identity

$$\{\det[I - ZA]\}^{-1} = 1 + \sum_{i=1}^{d} z_i a_{ii} + \frac{1}{2!}\sum\sum_{i,j=1}^{d} z_i z_j {}^{+}\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}^{+} + \cdots \tag{1}$$

where $Z = \operatorname{diag}(z_1 \dots z_d)$, $A = \{a_{ij}\}$ is square $d \times d$, the plus signs ${}^{+}|A|^{+}$ signify permanents, and the series continues indefinitely and converges provided the modulus of the largest eigenvalue of ZA is less than unity.

The purpose of the present note is to develop a similar expansion for arbitrary positive or negative powers of the characteristic polynomial $\det[I - ZA]$. It will be shown that the coefficients can again be expressed as the sums of certain matrix functions over all symmetrically placed submatrices built up from the elements of the original matrix. Specifically, we make the following definition, where

$$a\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = a_{i_1 j_1}, a_{i_2 j_2}, \dots, a_{i_k j_k}.$$

655 Avenue of the Americas, New York, NY 10010 0024-3795/88/$3.50

Definition. The α-permanent formed from a $k \times k$ matrix $A = \{a_{ij}\}$ is the matrix function

$$ {}^{\alpha}|A|^{\alpha} = \sum_{\sigma} \alpha^{m(\sigma)} a \begin{pmatrix} 1 & 2 & \cdots & k \\ \sigma(1) & \sigma(2) & \cdots & \sigma(k) \end{pmatrix}, \tag{2} $$

where the summation is extended over all permutations

$$ \sigma = \{\sigma(1) \quad \cdots \quad \sigma(k)\} $$

of the set of indices $(1,2,\ldots,k)$, and $m(\sigma)$ denotes the number of distinct irreducible cycles of the permutation σ.

The main result can now be stated as follows.

Theorem. *For all real α, and all $d \times d$ square matrices A,*

$$ \{\det[I - ZA]\}^{-\alpha} = 1 + \sum_{i=1}^{d} z_i(\alpha a_{ii}) + \frac{1}{2!} \sum\sum_{i,j=1}^{d} z_i z_j \, {}^{\alpha}\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}^{\alpha} + \cdots, \tag{3} $$

the series converging whenever the modulus of the maximum eigenvalue of ZA is less than unity.

The case $\alpha = 1$ corresponds to the permanent expansion, and the case $\alpha = -1$ to the usual expansion of the characteristic polynomial in determinants. In this case, as whenever α is a negative integer, the expansion is finite, i.e. is a polynomial in the z_i. In all other cases the expansion continues indefinitely.

If A is diagonal, the left hand side (LHS) of (3) reduces to a product of negative binomial terms, each with index α, and the expansion is obvious. This remark, together with the application of any similarity transformation which reduces ZA to diagonal or more generally triangular form, also serves to establish the criterion for convergence.

2. PROOF

To verify the general form of the coefficients, it is convenient to set $Z = zI$ (as may be done without loss of generality). We then take as starting point the identity

$$\{\det[I - zA]\}^{-\alpha} = \exp\{-\alpha \log\det[I - zA]\}$$
$$= \exp\left\{\alpha \sum_{1}^{\infty} \frac{z^k}{k} \operatorname{tr} A^k\right\}. \tag{4}$$

Then, following, e.g. Andrews [1], we express the exponentiated series in terms of partitions, where in general from

$$h(z) = \sum_{1}^{\infty} \frac{h_k z^k}{k!},$$

we obtain

$$\exp[h(z)] = 1 + \sum_{k=1}^{\infty} \frac{z^k}{k!} \sum_{\tau \in \mathscr{P}_k} \prod_{r=1}^{m(\tau)} h_{s_r}. \tag{5}$$

Here $\mathscr{P}_k$ denotes the family of all distinguishable partitions of the set $\{1,2,\ldots,k\}$, τ is a given member of $\mathscr{P}_k$, $m(\tau)$ is the number of distinct subsets, say $\mathscr{S}_1,\ldots,\mathscr{S}_{m(\tau)}$, of τ, and s_r is the cardinality of $\mathscr{S}_r$, $r = 1,\ldots,m(\tau)$, $s_1 + s_2 \ldots + s_{m(\tau)} = k$.

In our application the coefficients h_k are given by

$$h_k = \alpha(k-1)!\operatorname{tr}(A^k),$$

so that to establish the theorem we have to show that

$$\sum_{i_1,\ldots,i_k=1}^{d} \begin{vmatrix} a_{i_1 i_1} & \cdots & a_{i_1 i_k} \\ \vdots & & \vdots \\ a_{i_k i_1} & \cdots & a_{i_k i_k} \end{vmatrix}^{\alpha} = \sum_{\tau \in \mathscr{P}_k} \alpha^{m(\tau)} \prod_{r=1}^{m(\tau)} (s_r - 1)!\operatorname{tr}(A^{s_r}). \tag{6}$$

The LHS of this expression may be regarded as a double sum, first over the ordered sets of indices, and second over the permutations of those indices within the α-permanents. Let us reverse the order of these sums, at the same time classifying the permutations which occur, first by the partition $\tau = \tau(\sigma)$ which determines the number and choice of the elements forming the distinct irreducible cycles of σ [note $m(\sigma) = m(\tau)$] and then by the particular cycle constructed from the elements in the given subset of indices. If a subset of the partition, say $\mathscr{S}$, contains $s = |\mathscr{S}|$ elements, then $(s-1)!$ distinct irreducible cycles may be constructed from those elements, as the first we conventionally hold fixed and the rest may be permuted among themselves. For each subset $\mathscr{S}_r$ of τ, let $\mathscr{C}_r$ denote the family of all irreducible cycles built from the elements of $\mathscr{S}_r$, σ_r denote a general element of $\mathscr{C}_r$, and $s_r = |\mathscr{S}_r|$. With this notation we may, after some rearrangement, rewrite the LHS of (6) in the form

$$\sum_{\tau \in \mathscr{P}_k} \prod_{r=1}^{m(\tau)} \left\{ \sum_{\sigma_r \in \mathscr{C}_r} \sum_{i_1, \ldots, i_{s_r} = 1}^{d} a \begin{pmatrix} i_1 & \cdots & i_{s_r} \\ \sigma_r(i_1) & \cdots & \sigma_r(i_{s_r}) \end{pmatrix} \right\}. \tag{7}$$

Thus to complete the proof it will be enough to show that for every $s = 1, 2, \ldots$

$$\sum_{\sigma \in \mathscr{C}_r} \sum_{i_1, \ldots, i_s = 1}^{d} a \begin{pmatrix} i_1 & \cdots & i_s \\ \sigma(i_1) & \cdots & \sigma(i_s) \end{pmatrix} = (s-1)!\, \mathrm{tr}(A^s). \tag{8}$$

To this end consider any particular term in the LH sum, say $a_{11}a_{12}a_{22}a_{21}$ in the case $s = 4$. We may characterize any such term by the set of initial indices, here 1-1-2-2, which, when acted on by the canonical cycle (1234), produce the second set of indices (1-2-2-1) defining the given term. We ask how often this term occurs, first as we vary the cycle acting on the initial set of indices, and then as we vary the indices themselves.

If the initial sequence contains j distinct indices, repeatedly respectively $r_1, r_2, \ldots, r_j$ times $(r_1 + r_2 + \cdots r_j = s)$, the same second sequence will be produced by any cycle which merely permutes identical indices, there being $(r_1 - 1)! r_2! \cdots r_j!$ such cycles (recall the first element remains conventionally fixed). The same product will also recur, albeit with a different ordering of the terms, if a cycle is applied which has the effect of interchanging the roles of the fixed first index and any one of its later appearances. For example, if we interchange the role of the two 1's in the sequence 1-1-2-2, we are led to the cycle (1342) in place of the basic cycle (1234), and can check that (1342) and (1432), as well as (1234) and (1243) all give rise to the term $a_{11}a_{12}a_{22}a_{21}$. This process fails if the initial sequence of indices already contains repeat cycles, for then interchanging the roles of the first index and its appearance at the beginning of any of the repeat cycles merely reproduces cycles already considered. Thus for example the term $a_{12}^2 a_{21}^2$, corresponding to the initial set

(1-2-1-2), is produced by the cycles (1234) and (1432), the latter arising both when we permute the 2's and when we interchange the 1's.

Thus in general the total number of cycles reproducing the given term when applied to the initial set of indices is $r_1!, r_2!, \ldots, r_j!/t$, where t is the number of repeat cycles.

If now we consider the effect of varying the indices, it will be seen that exactly the same number of cycles will be obtained from any initial set containing the same number of repetitions of the individual indices as the original set. The number of such sets is $s!/(r_1!, \ldots, r_j!)$. Thus the overall number of repetitions of the given term is just $s!/t$, t being the number of repeat cycles in the original specification of the indices.

On the trace side also, any term such as $a_{11}a_{12}a_{22}a_{21}$ can be characterized by the set 1-1-2-2 of first indices, and will recur once for each choice of index as first term (i.e. s times in all), save in the presence of repeat cycles, when the original terms are repeated with each cycle and the number of occurrences is reduced from s to s/t. Thus 1-2-1-2 produces only two terms from the trace, viz. $a_{12}a_{21}a_{12}a_{21}$ and $a_{21}a_{12}a_{21}a_{12}$, whereas 1-1-2-2-arises four times, as $a_{11}a_{12}a_{22}a_{21}$, $a_{12}a_{22}a_{21}a_{11}$, $a_{22}a_{21}a_{11}a_{12}$, and $a_{21}a_{11}a_{12}a_{22}$. Overall, the number of occurrences in the RHS of (8) is therefore $(s-1)! \times s/t = s!/t$, the same as for the LHS.

This completes the proof.

3. DISCUSSION

The expansion in the theorem, being linear in powers of α as well as in powers of z, can be extended to any function of the characteristic polynomial which can be expressed as a linear combination of positive or negative powers of its argument. In particular, if F is a probability distribution for α,

$$\int \{\det[I - ZA]\}^{-\alpha} dF(\alpha)$$

$$= 1 + \sum z_i \mu_1 a_{ii} + \frac{1}{2!} \sum \sum z_i z_j \,{}^{(F)}\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}^{(F)} + \cdots,$$

where

$${}^{(F)}\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}^{(F)} = \sum_{\sigma} \mu_{m(\sigma)} a\begin{pmatrix} i_1 & \cdots & i_k \\ \sigma(i_1) & \cdots & \sigma(i_k) \end{pmatrix}$$

and $\mu_m = \int \alpha^m dF(\alpha)$, the sum being taken over all permutations σ as in (2). Here the series will converge provided the maximum eigenvalue of ZA has modulus less than the radius of convergence of the series $\sum \mu_k z^k$.

Similarly one may consider analytic or meromorphic functions of the characteristic polynomial.

The expansions have applications, in particular, to multivariate binomial and negative binomial distributions, whose probability generating functions are proportional to expressions of the form $\{\det[I - ZA]\}^{\pm \alpha}$. For example, the class of multivariate negative binomial distributions considered recently by Griffiths and Milne [3] is of this form. Further details and other probability applications are given in [5].

A particularly lucid account of the original expansion (1) in the context of multilinear algebra is given in [2]. We refer to this paper and our original note [4] for further background material.

The proof of the main theorem is an adaptation of an argument originally suggested to me by R. B. Davies. The present paper had its origins in a seminar I gave at Cambridge while visiting England on a Royal Society grant. I would like to express my gratitude to the individuals and institutions concerned for their help and support.

REFERENCES

1 G. Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, Mass., 1976.
2 A. Blokhuis and J. J. Seidel, An introduction to multilinear algebra and some of its applications, *Philips J. Res.* 39:111–120 (1984).
3 R. C. Griffiths and R. K. Milne, A class of infinitely divisible multivariate negative binomial distributions, *J. Multivariate Anal.*, 22:13–23 (1987).
4 D. Vere-Jones, An identity involving permanents, *Linear Algebra Appl.* 63:267–270 (1984).
5 D. Vere-Jones, α-permanents and their application to a class of multivariate negative binomial and binomial distributions, preprint, Victoria Univ., 1986.

Received 17 December 1987; final manuscript accepted 11 February 1988